*Am. J. Hum. Genet.* 60:255–262, 1997

# 1996 WILLIAM ALLAN AWARD ADDRESS
# Algorithms and Inferences: The Challenge of Multifactorial Diseases

Robert C. Elston

Department of Epidemiology and Biostatistics, Rammelkamp Center for Education and Research, MetroHealth Campus, Case Western Reserve University, Cleveland

On learning that I was to receive the Allan Award for 1996, I had mixed feelings of elation, dismay, and surprise. Elation, because receiving the Allan Award is an honor bestowed on so few; dismay, because I realized that it would mean I would have to face this audience with something interesting to say; and surprise, because the decision to give me the award must have been made shortly after Kevin Davis, the editor of *Nature Genetics,* published the following statement regarding a letter I had just coauthored: "We disagree with this approach for three reasons. It is intellectually wrong. . . . It is historically inconsistent. . . . It is technically antediluvian" (Lander and Kruglyak 1996, p. 358).

The letter I had coauthored concerned the "genetic dissection of complex traits," a subject that has in recent years received a surge of interest. I felt certain that the Awards Committee would want me to give an address on a topic of current interest, and what could be of more current interest in the area I am known to work in? And why would my views on this topic be of any interest if they are intellectually wrong, historically inconsistent, and technically antediluvian? So what *should* I talk about?

My mind went back ~20 years to University Day at the University of North Carolina at Chapel Hill. The guest speaker was the comedian Andy Griffith, an alumnus of the University. After being introduced, he walked up to the podium and looked around at the audience all in full academic regalia. He then quickly told three or four jokes. "There," he said, "tha's what ah do for a living." Well, figure 1 illustrates what I, an animal-breeder-turned-biostatistician, do for a living. Figure 1*a* illustrates the main results of the Elston-Stewart algorithm, while figure 1*b* shows the Haseman-Elston method. There must be many persons in the audience for whom this is Greek; and I suspect there are even some who would find it more comprehensible if it were Greek. My friend and one-time colleague Mary K. Pelias calls it, perhaps appropriately, "chicken scratches."

I have entitled my talk today "Algorithms and Inferences" because I want to show how, when summary statistics are calculated from a set of observations, the conclusions that are drawn depend on both the algorithms used to calculate them and the rules of inference that one believes are appropriate. This is especially true in the case of the genetic dissection of multifactorial diseases, for which we now have very powerful tools, including sophisticated computer programs. Whereas a sharp scalpel in the hands of a skillful surgeon can be used for healing, the same scalpel in the hands of an unskilled person can cause harm. Similarly, whereas the powerful computer programs currently available can, in the hands of a skillful genetic analyst, be used to advance our genetic knowledge, the same programs in the hands of an unskilled person can cause harm by leading to invalid conclusions. I shall first briefly review how some of the algorithms work and then give an overview of some principles of inference. Finally, I shall discuss problems that arise in the search for individual genes underlying complex diseases. I shall confine myself to the subject of linkage analysis, i.e., the search for familial cosegregation of genes underlying disease with genetic markers. Suppose a person receives from one parent a chromosome bearing a D allele at a disease locus and an M allele at a marker locus but, from the other parent, the alleles d and m; then figure 2 summarizes what may result when such a person forms gametes to be passed on to the next generation. Each gamete may be a nonrecombinant or a recombinant. The proportion of gametes that are expected to be recombinants is the recombination fraction, a parameter that is a measure of the genetic distance between two loci and is usually denoted $\theta$. If $\theta < .5$ there is linkage between the two loci; whereas if $\theta = .5$ there is no linkage.

## Algorithms

Computer algorithms are used to calculate statistics. It is common to think of the term "algorithm" as referring only to the method by which a particular result, in our case a statistic, is calculated. More important,

Received November 25, 1996. This address was delivered November 2, 1996, in San Francisco.

Address for correspondence and reprints: Dr. Robert C. Elston, Department of Epidemiology and Biostatistics, MetroHealth Medical Center, 2500 MetroHealth Drive, Room R259, Cleveland, OH 44109-1998.


0002-9297/97/6002-0002$02.00

(a)

The likelihood of the entire data can be expressed as the sequence of operations

$$\Gamma_0(\Gamma_1(\Gamma_2(\Gamma_3\ldots)))$$

where $\Gamma_j = \prod_{i_j} \sum_{s_j=1}^{k} p_{s_{j-1}t_{j-1}s_j} g_{s_j}(x_{i_0 i_1 \ldots i_j}) \sum_{t_j=1}^{k} \psi_{t_j} g_{t_j}(y_{i_0 i_1 \ldots i_j})$,

provided we define $p_{s_{j-1}t_{j-1}s_j}$ as $\psi_j$ when $j = 0$.

(b)

$$E(Y_{jt} \mid \hat{\pi}_{jm} = 0) = \sigma_e^2 + 2\Psi\sigma_g^2$$

$$E(Y_{jt} \mid \hat{\pi}_{jm} = \tfrac{1}{2}) = \sigma_e^2 + \sigma_g^2$$

$$E(Y_{jt} \mid \hat{\pi}_{jm} = 1) = \sigma_e^2 + 2(1-\Psi)\sigma_g^2$$

$$E(Y_{jt} \mid \hat{\pi}_{jm}) = [\sigma_e^2 + 2\Psi\sigma_g^2] + 2(1-2\Psi)\sigma_g^2\hat{\pi}_{jm}$$

$$= [\sigma_e^2 + 2(1-2c+2c^2)\sigma_g^2] - 2(1-2c)^2\sigma_g^2\hat{\pi}_{jm}$$

**Figure 1** Statistical basis underlying two methods of linkage analysis. *a*, Elston-Stewart algorithm (Elston and Stewart 1971). *b*, Haseman-Elston regression method (Haseman and Elston 1972).

however, is just what is being calculated, and I use the term to refer to that also. We are often interested in estimating the recombination fraction θ from a set of families segregating for a disease, with each family member being typed at a marker locus. If we know the recombination fraction, as well as the mode of inheritance of the disease and marker, we can calculate the probability of any particular data outcome. This is expressed as a *probability function*, $P(\text{data} \mid \text{recombination fraction})$. For example, to take a simple situation, suppose we have a mating in which there are $n$ offspring, and, after typing all the offspring, it is possible to unequivocally score each as a recombinant or not. If we let $r$ be the number of recombinants, then this probability function is $P(r \mid \theta) = \binom{n}{r}\theta^r(1-\theta)^{n-r}$. Substituting $r = 0, 1, 2, \ldots,$ or $n$ into this expression gives us the appropriate probability for any value of $r$. Here we consider θ to be a known fixed value, and the value of the function depends on the variable $r$, the number of recombinants that will be observed.

Once we have typed the offspring, we know the outcome $r$, so this is now no longer a variable. Sir Ronald Fisher, one-time Balfour Professor of Genetics at Cambridge University, defined the *likelihood function* of a parameter θ for a particular set of data $r$, L(recombination fraction | data) to be mathematically the same as the probability function, but now we consider $r$ to be fixed, and the different values that the function can take on depend on the different possible values of the unknown parameter θ. Thus the two expressions are the same, the only difference being that the likelihood L is considered as a function of the recombination fraction θ, while the probability $P$ is considered as a function of the data outcome $r$. (It so happens that it is only the *relative* magnitudes of the likelihood $L(\theta \mid r)$ that are of interest, so that the strict definition is one of proportionality instead of one of equality: $L(\theta \mid r) \propto P(r \mid \theta)$.) The relative magnitudes of the likelihood for different possible values of θ can be used to choose among them, and the maximum-likelihood estimate of θ is that value of θ that maximizes the likelihood, i.e., the value that makes it most probable for our sample data to have occurred. Fisher showed that maximum-likelihood estimates have desirable properties in large samples. It is therefore of interest to calculate likelihoods for large pedigrees, which can be very informative for the linkage analysis of a rare monogenic disease.

The purpose of the Elston-Stewart algorithm was to calculate the likelihood for a large pedigree under a broad class of genetic models, one of which involved linkage between a marker and any monogenic trait, whether a discrete disease trait or a continuous quantitative trait. The algorithm caused concern initially, because the calculations start at the bottom of the pedigree instead of following the flow of genes down through the pedigree (Ott 1985, p. 33). However, it so happens that the proposed algorithm is very efficient in terms of computing time. The principle that brings about this efficiency can be seen in a simple example. Figure 1 shows that the Elston-Stewart algorithm involves summations (Σ) and products (Π). Now the order in which these operations are performed can affect how much computing time is required. For example, the calculation $(2\times3)+(2\times4)+(2\times5) = 24$ requires three multiplications and two additions, but changing the sequence of operations to $2\times(3+4+5) = 24$ requires only one multi-

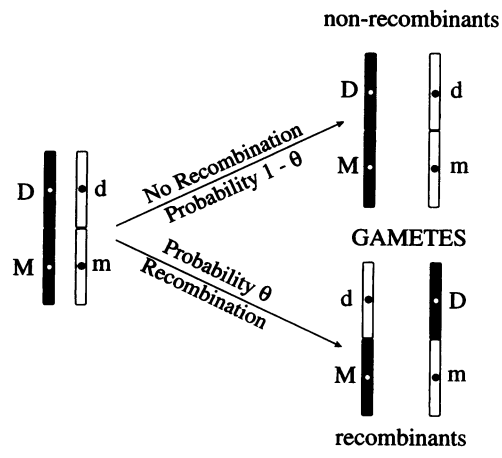


**Figure 2** Formation of nonrecombinant and recombinant chromosome segments (*right*), from the corresponding parental segments (*left*).

plication and two additions. Although the savings in this simple example are small, multiple applications of the same principle can make the difference between a single likelihood calculation taking $\leq 1$ s, versus $\geq 1$ h. The Elston-Stewart algorithm ensures that the time to compute a likelihood increases only linearly, rather than exponentially, with the number of persons in the pedigree. On the other hand, the time required increases exponentially as the number of loci increases, so that the method soon becomes infeasible if we try to analyze a disease and multiple marker loci simultaneously, i.e., multipoint linkage analysis. For this purpose Lander and Green (1987) proposed an alternative algorithm whose computation time increases linearly with the number of loci but exponentially with the number of pedigree members.

In both these algorithms many multiplications of probabilities are performed. Now, when two events are independent, their joint probability is obtained by multiplying together their individual probabilities. Conversely, if the joint probability is the product of the individual probabilities, then the events are independent. The fact that in pedigree likelihoods we find many probabilities being multiplied together implies that many assumptions are being made about certain things being independent. It is these very assumptions (in addition to the specific sequence in which the additions and multiplications are performed) that make the likelihood computations feasible, and they need to be carefully scrutinized. Both the Elston-Stewart and Lander-Green algorithms have been substantially improved since they were first proposed. It is important to differentiate those improvements that are computational in nature, such as the sequence in which operations are performed, from those that become possible by making restrictive assumptions that may not always hold.

The approach to linkage analysis proposed by Haseman and Elston (1972) had a different purpose. Whereas the Elston-Stewart approach was intended specifically for the analysis of well-defined monogenic traits, albeit of multifactorial etiology because of environmental influences, the Haseman-Elston approach was put forward as a method to detect linkage for traits of ill-defined etiology, such as psychological traits and mental diseases. It is easiest to understand the principle underlying it in the case of a continuous trait, but the method can be used for any quantitative trait, including a disease trait which can be quantified, e.g., by letting 0 = absence of disease and 1 = presence of disease. (In fact, the first published application of the method was to a disease trait [Elston et al. 1973].) Consider a sample of sib pairs who have been typed for a polymorphic marker and on whom a trait has been measured. At any marker locus, the sib pairs must share zero, one, or two alleles identical by descent (IBD) (i.e., direct copies of the same parental

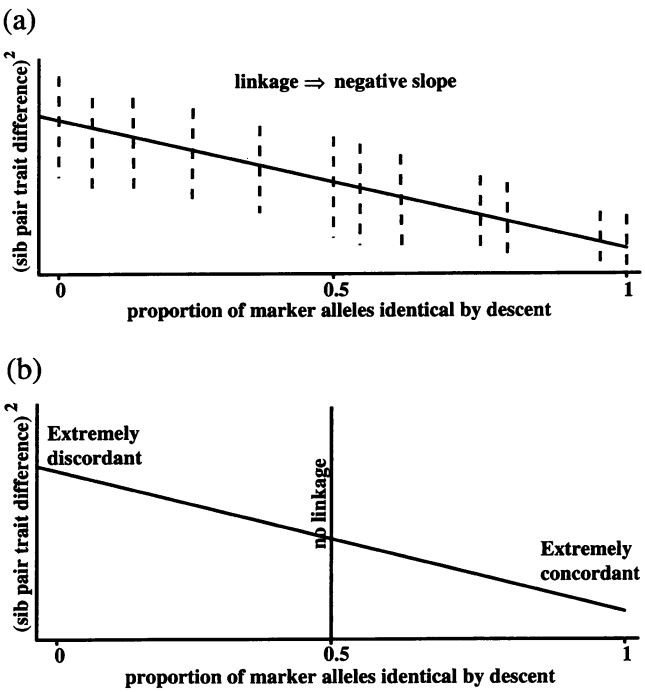


**Figure 3** Haseman-Elston (1972) method. *a*, Plot of data points illustrating the regression shown in *b*: (sib-pair trait difference)$^2$ = $Y$, linkage corresponds to $.5 < \Psi \leq 1$, in fig. 1*b*. *b*, Illustration of extremely discordant and extremely concordant pairs being the most powerful to detect linkage.

allele), and, if there is linkage to a locus underlying the trait of interest, they will correspondingly tend to share zero, one, or two alleles IBD at that trait locus. The result is that, if we plot the squared sib-pair trait difference against the proportion of alleles shared IBD at the marker locus, the points will show the trend of a line with a negative slope, as illustrated in figure 3*a*. Sibs who are alike in the trait will tend to be alike at a linked marker, and sibs who differ in the trait will tend to differ at a linked marker. If there is no linkage, such a trend will be absent. In many instances it may not be possible to determine exactly the proportion of marker alleles shared (which must be 0, .5, or 1), but, with a knowledge of the marker-allele frequencies, this may be estimated as a weighted average, e.g., of 0 and .5 or of .5 and 1, as can be seen in figure 3*a*. We can fit a line to the data points by "least squares," i.e., by minimizing the average squared distance of the points to the line, and so determine if the slope is negative. Large-sample theory can be used to determine whether a negative slope is likely to reflect the true situation in the population sampled or is a chance occurrence that would disappear if more sib pairs were studied. I call this method "model free," in that it does not model the mode of inheritance of the trait being studied and its validity does not depend on any such assumptions, as opposed to a "model-based"

analysis that models a particular mode of inheritance. The method has also been called a "robust" or "nonparametric" method, but the former term can describe a model-based analysis performed using the Elston-Stewart algorithm, and model-free methods can be parameterized in terms of the proportion of marker alleles shared IBD, lending them amenable to likelihood analysis (Risch 1990). Although it has been stated that these model-free methods assume a genetic model for the trait being studied (Blangero 1993; Whittemore 1996), this is not the case: their *validity* does not depend on any such assumptions. This so-called sib-pair method has been extended in many ways, and in particular we can ask if the proportion of alleles shared IBD is significantly different from .5. It can be seen in figure 3*b* that the most powerful samples of sib pairs to detect departure from this null hypothesis are ones in which the sib pairs are either extremely discordant or extremely concordant, those around the middle of the line being of little use (Risch and Zhang 1996).

## Inferences

Statistical significance is another concept due to Fisher. If we wish to judge to what extent a set of data supports the hypothesis of linkage, we can ask how probable, if in fact the hypothesis of no linkage (the null hypothesis) is true, is it for the data—or any other outcome even more suggestive of linkage—to have occurred. If this probability is small, then we have what Fisher called a "logical disjunction"—either the null hypothesis is not true (we do have linkage), or we have observed a rare event. The rarity of such an event, called "the *P* value" or "empirical significance level," is a measure of how unlikely the data are if there is no linkage. Thus, the smaller the *P* value, the more we are inclined to believe that there is linkage. Performing a significance test is the same as determining a *P* value, and we say the result is significant at the *P* level. But it must be realized that the calculation of a *P* value involves many more assumptions than just the absence of linkage; a probability model is assumed, and it is significant departure from the totality of this model that is being evaluated.

It is, to my mind, unfortunate that emphasis is sometimes placed more on hypothesis testing than on significance testing. In hypothesis testing, rather than reporting a quantitative measure that reflects how unlikely the data are if there is no linkage, we set up a strict algorithm that tells us whether to declare that linkage has been found or not, with the understanding that, if, in fact, there is no linkage (more correctly, if the null hypothesis is true), the probability that we should declare linkage is limited to some small level $\alpha$, the probability of what is called the "type I error." Of course, one also makes an error if one declares there is no linkage when, in fact, there is linkage. In a well-designed study, the probability of this type II error, denoted $\beta$, is also controlled. Unfortunately, this cannot be done with any precision when studying complex diseases (reviewers of grant applications, please note!), because the probability of type II error can only be calculated if one specifies, in addition to what is needed to control type I error, unknown details about the mode of inheritance of the disease one is studying.

**Table 1**

**Classification of Linkage Tests by Whether Linkage Is Present and Whether Linkage Is Declared**

| | True State of Nature | | |
|---|---|---|---|
| Declare | No Linkage | Linkage | Total |
| No linkage | $a$ | $b$ | $a + b$ |
| Linkage | $c$ | $d$ | $c + d$ |
| Total | $a + c$ | $b + d$ | $n$ |

Hypothesis testing, with its rigorous algorithm for coming to a clear-cut answer—the null hypothesis is declared either true or false—has great utility for those situations in which a decision must be made, such as in business situations and in medical practice. It is beloved by editors and reviewers, because it makes their lives easy. I once had a reviewer tell me that it was wrong to report a result as being significant at the 6% level because that was not significant! But hypothesis testing is unscientific. In science we have the luxury of not needing to make snap decisions. It is recognized that theories—and even laws—may at any time fall by the wayside. Mendel's second law, which in essence said that linkage does not exist, is one example that all geneticists should remember. Even significance testing, which at least tries to quantify our uncertainty, has its limitations. What is of scientific interest is not so much the probability of making a mistake *when the null hypothesis is true* as the probability of making a mistake *when declaring that the null hypothesis is false*. If we consider a large number of tests for linkage classified by whether or not there is in fact linkage and whether or not linkage is declared, the numbers in the fourfold table being *a, b, c* and *d* (table 1), the usual probability that is controlled is the probability of type I error when there is no linkage, $c/(a + c)$: of greater interest is the *posterior* probability of type I error when linkage is declared on the basis of a test, $c/(c + d)$. More generally, we should really like to know, at the end of study, the probability that we have found a linkage, as pointed out by Cedric Smith (1959) >35 years ago.

Consider a large population of conceivable linkage

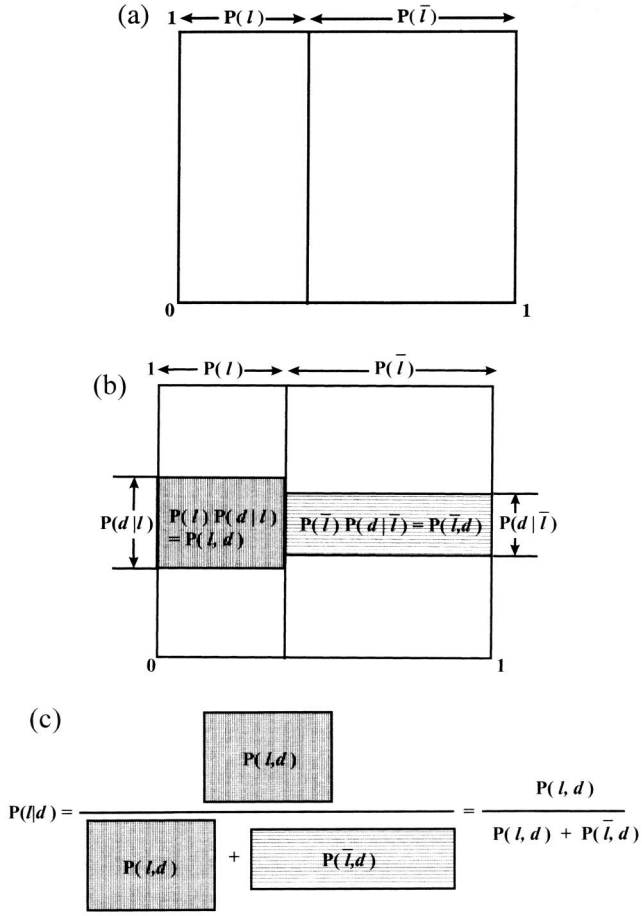


**Figure 4** Derivation of the posterior probability of linkage assuming the simple dichotomy "linkage" vs. "no linkage." See text for details.

tests in which the probability of there actually being linkage is $P(l)$ and the probability of there not being linkage is $P(\bar{l}) = 1 - P(l)$, as illustrated in figure 4*a*. We observe a set of data $d$ and can calculate the probability of such an observation (i) given that there is linkage and (ii) given that there is no linkage: these are the vertical distances indicated in figure 4*b*. The left hatched rectangle in figure 4*b* represents the joint probability of linkage and our data, while the right hatched rectangle represents the joint probability of no linkage and our data. Given our data, these two hatched areas are the only areas that are relevant, so that the probability of linkage, given our data, is as shown in figure 4*c*, which is an example of Bayes's theorem. Recall now that linkage corresponds to any value of $\theta < .5$ and that the probability of our data depends on $\theta$. A more accurate representation is therefore that given in figure 5*a*, with the resulting probability of linkage being as indicated in figure 5*b*. All this assumes that we know, before we start our study, the *prior* probability distribution of $\theta$; i.e., we need to know in what proportion of all our conceivable linkage tests there will in fact be linkage, and, if there is, what are the corresponding probabilities of the various possible values of the recombination fraction $\theta$.

I mentioned earlier that it is only the relative magnitudes of likelihoods that are of interest. Suppose we are interested in two particular hypotheses—say, two values of the parameter $\theta$: $\theta = \theta'$ and $\theta = .5$. For this situation, George Barnard (1949) defined the term "lod" to be the logarithm of the "odds" of $\theta = \theta'$ versus $\theta = .5$. However, he differentiated forward odds from backward odds. The odds of one event versus another is defined to be the probability of the first event divided by the probability of the second. If we think of linkage as an event of interest after we have gathered a set of data, then the lod is the logarithm of the forward odds for $\theta'$ versus .5, which is what we would like to know in order to choose between the two values of $\theta$. But, without knowledge of the prior distribution of $\theta$, only the backward lod can be calculated (fig. 6). What is meant by lod in linkage analysis is this *backward* lod, the logarithm of the ratio of two likelihoods, which is quite a different thing. One still occasionally sees in the linkage literature a (backward) lod interpreted as the logarithm of the odds for linkage, but this is incorrect unless one specifies that the *backward* odds is being referred to.

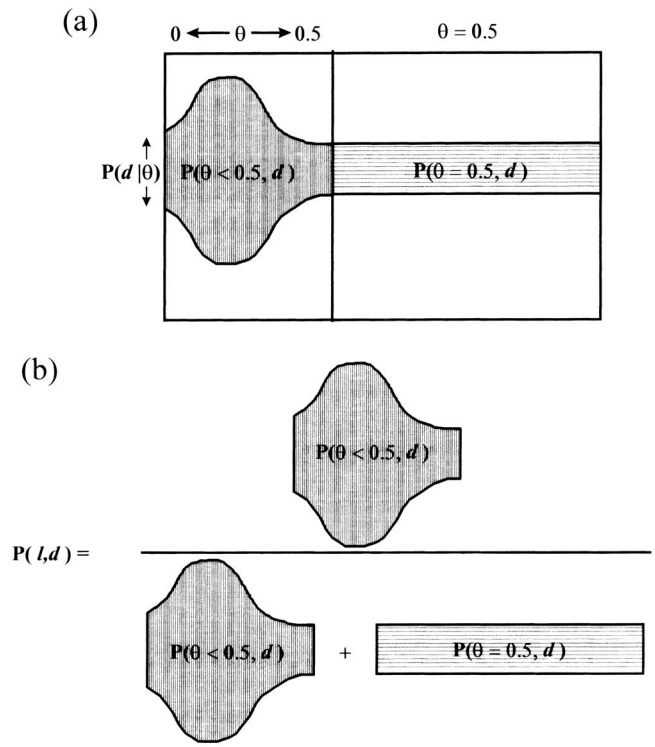


**Figure 5** Derivation of the posterior probability of linkage taking account of the fact that linkage corresponds to a recombination fraction ($\theta$) between 0 and .5.

Logarithm of the forward odds

$$= \log \frac{P(\theta = \theta' | d)}{P(\theta = 0.5 | d)}$$

= log of the odds for $\theta = \theta'$ versus $\theta = 0.5$

Logarithm of the backward odds

$$= \log \frac{P(d | \theta = \theta')}{P(d | \theta = 0.5)} = \log \frac{L(\theta = \theta' | d)}{L(\theta = 0.5 | d)}$$

= log of the likelihood ratio for $\theta = \theta'$ versus $\theta = 0.5$

**Figure 6** Distinction between the logarithm of forward odds and the logarithm of backward odds. See text for details.

Newton Morton (1955), the first recipient of the Allan Award, introduced lods into linkage analysis. He proposed taking logarithms to base 10, so that a lod of 1, 2, or 3 would correspond, respectively, to a value of $\theta$ being 10, 100, or 1,000 times as "likely" as no linkage—i.e., the observed data being 10, 100, or 1,000 times as probable, given that value of $\theta$ rather than $\theta = .5$. He considered the problem of detecting linkage between a marker and a well-defined monogenic disease, on the assumption that the disease locus has equal probability of being located anywhere on the genetic map. With this prior uniform distribution, he was able to come up with an approximate posterior probability of making an error when declaring linkage. He recommended that, to keep this probability down to ~5%, linkage should only be declared significant when a lod of 3 is reached. It can be shown that this always corresponds to a $P$ value of $<10^{-3}$ and that, for large samples, it corresponds to a $P$ value of $\sim 10^{-4}$.

Gustave Malécot (1947), using linkage analysis as an example, showed that choosing a single cut point for the probability of type I error, in order to choose whether or not to believe the null hypothesis, is irrational. Much as though all scientists want to be objective, he pointed out that scientific knowledge must be subjective. Whatever experimental result is found, any rational scientific interpretation must necessarily depend on the prior beliefs of the interpreter. This principle must apply with even greater force in the study of complex diseases, to which I now turn.

## Complex Diseases

For the purposes of genetic analysis, we can take as an operational definition of a complex disease any disease for which we have no reasonable basis for determining the prior probability of linkage, because we do not know with certainty how many—or even if—genes are involved in its etiology. It is therefore impossible to apply Bayes's theorem to arrive at a posterior probability of linkage. Furthermore, we either have to use a model-free method of analysis, or, if we use a model-based method of analysis, we need to consider the appropriateness of the assumptions we are making. The best way to determine the critical assumptions is by examining the particular likelihood function being used. Typical assumptions are that (*a*) segregation at a single locus is the sole cause of familial aggregation of the trait being studied, (*b*) all relationships among family members are known without error, (*c*) there is no pleiotropic effect of the marker locus on the trait (whether epistatic or not), (*d*) the only unknown parameters are recombination fractions, and (*e*) (if any founder marker genotypes are unknown) the marker genotype frequencies are known without error. It is noteworthy that, whereas individually these assumptions may not be necessary to obtain a valid $P$ value, jointly they can be critical. Thus, when a small $P$ value is calculated, it may be one of these assumptions that is being evaluated, in addition to absence of linkage. For this reason the linkage analysis of a complex trait is fraught with difficulty.

Some have argued that the best way to unravel the genetic etiology of a complex disease is by studying candidate genes, guessing what biochemical pathways are involved. Others have argued that we should perform a global search of the whole genome, using roughly equally spaced markers as a dragnet to search for the approximate location of genes, by linkage. Which is better depends on many factors, not least among them being the prior probability that one's guesses regarding candidate genes are correct. I believe that both approaches will be necessary for some time to come. Either way, many markers will be tested, either biochemical markers that may be functionally related to the disease or nonfunctional genetic markers that are indicators of gene location. Statisticians have long studied the "multiple comparisons" problem of interpreting $P$ values when many tests are performed. If we say that the result of a test is significant when the $P$ value is less than, say, .05, then we can increase our chances of finding a significant result, even if the null hypothesis is true, by increasing the number of tests performed. Just by chance alone, the smallest of 1,000 $P$ values will be smaller than the smallest of 100 $P$ values.

Because we now have so many markers available for scanning the whole genome by linkage analysis, we need to allow for this when interpreting the smallest $P$ values from such a scan. In a thoughtful article, Lander and Kruglyak (1995) discussed this problem in some detail. Among other things, they proposed that all $P$ values found in linkage analysis should be adjusted to allow for the large number of tests in a dense scan of the whole genome, whether or not such a large number of tests was actually performed, with the result that a LOD of $\geq 3.3$ would be required before declaring linkage sig-

**Table 2**

**Comparison of Hypothetical Results from Two Different Sib-Pair Studies Looking at a Complex Genetic Trait**

| Study | No. of Pairs | *P* Value[a] | Importance[b] |
|---|---|---|---|
| A | 200 | $1.99 \times 10^{-5}$ | Significant linkage |
| B | 100 | $2.01 \times 10^{-5}$ | Suggestive linkage |

[a] For linkage.
[b] Based on the Lander and Kruglyak (1996) classification scheme.

nificant. In a letter that applauded their efforts, Witte, Elston, and Schork (Witte et al. 1996) pointed out that this particular proposal does not have a sound statistical basis. We also pointed out that blanket use of the cut points they proposed, without any thought being given to other aspects of the particular study, could have unfortunate consequences. For example, table 2 shows what could be the results of two sib-pair linkage studies. With virtually identical *P* values, the first would be classified as "significant linkage" under their proposed scheme, while the second would be classified as "suggestive linkage." Furthermore, the second study could well be the more important result, because it is based on a smaller sample size! We suggested that authors of linkage studies should report precise *P* values (not adjusting them for multiple comparisons that were never made), so that they can be properly interpreted by the reader.

This letter provoked a rejoinder, giving the three reasons why our suggestion was wrong (Lander and Kruglyak 1996, p. 358):

1. "It is intellectually wrong, in that it implies that investigators who happen to encounter chance fluctuations early in a project should accord them greater weight than those who find exactly the same evidence late in a project." Is it intellectually wrong to interpret the smallest of 1,000 *P* values differently from the smallest of 100 *P* values?
2. "It is historically inconsistent, in that it flies in the face of longstanding practice in human genetics of using the lod threshold of 3, which is premised on a whole-genome significance level." Was the lod threshold of 3 premised on a whole-genome significance level, or was it premised on the low prior probability of linkage to a single randomly placed marker, which can be calculated for a monogenic disease? If it is argued that this is the same thing, is it historically consistent to change the lod threshold of 3 to one of $\geq 3.3$, depending on the type of study?
3. "Finally, it is technically antediluvian, in that it ignores the fact that technology has progressed to the point where a whole-genome scan is fast becoming the quantum of publication." Do we want to deny anyone who has not completed a whole-genome scan the opportunity of publishing any results?

I leave it to you to decide what is the best approach to reporting results of analyzing complex diseases, though I have little doubt that editors will have the final word. In the past when I have been involved in correspondence to the editor of a scientific journal, it has been the norm for the editor to circulate each letter to all those concerned, to be sure that, prior to publication, all authors agree as to the facts; for then, discussion and differences of opinion will lead to more light than heat. The editor of *Nature Genetics* does not subscribe to this policy.

On a more humorous note, I should point out that our original letter tried to highlight what was perhaps meant by a proposal made by Lander and Kruglyak that "more informal vehicles" should publish less "significant" findings. The table included an additional column, headed "Acceptable Journal." For study A, we had entered *Nature* and for study B we had entered *Nature Genetics*. This drew the following response from Laura Goodman, the assistant editor of *Nature Genetics:* "Although we understand the point you are trying to make with this table, we do not wish to present what may be inaccurate representations in terms of what manuscripts are or are not appropriate for publication in both *Nature* and *Nature Genetics*. Simply removing the column 'Acceptable Journal' should still allow you to make the main point that, by having arbitrary cutoffs for the *P* values, minor differences in *P* values can constitute major differences in the manner in which this data is viewed with regard to its relative importance." (Letter to J. S. Witte). We understood that our example was perhaps inaccurate and perhaps even a little demeaning, so, in our revised version of the letter, we changed the last column to make it more palatable to *Nature Genetics*. We entered *Nature Genetics* for study A and *National Enquirer* for study B. As can be seen in table 2, the table was published without the last column—and without further comment.

In conclusion, I wish to thank all those, far too numerous to name individually, who have helped me one way or another to arrive at where I am today: family, friends, students, staff, colleagues, and teachers.

## Acknowledgments

This work was supported in part by U.S. Public Health Service research grant GM 28356 from the National Institute of General Medical Sciences and by resource grant P41 RR03655 from the Division of Research Resources.

## References

Barnard GA (1949) Statistical inference. J R Stat Soc [B] 11: 116–139

Blangero J (1995) Genetic analysis of a common oligogenic trait with quantitative correlates: summary of GAW9 results. Genet Epidemiol 12:689–706

Elston RC, Kringlen E, Namboodiri KK (1973) Possible linkage relationships between certain blood groups and schizophrenia or other psychoses. Behav Genet 3:101–106

Elston RC, Stewart J (1971) A general model for the genetic analysis of pedigree data. Hum Hered 21:523–542

Haseman JK, Elston RC (1972) The investigation of linkage between a quantitative trait and a marker locus. Behav Genet 2:3–19

Lander ES, Green PP (1987) Construction of multilocus genetic linkage maps in humans. Proc Natl Acad Sci USA 84: 2363–2367

Lander ES, Kruglyak L (1995) Genetic dissection of complex traits. Nat Genet 11:241–247

——— (1996) Genetic dissection of complex traits. Nat Genet 12:355–358

Malécot G (1947) Les critères statistiques et la subjectivité de la connaissance scientifique. Ann Univ Lyon 10:43–74

Morton NE (1955) Sequential tests for the detection of linkage. Am J Hum Genet 7:277–318

Ott J (1985) Analysis of human genetic linkage. Johns Hopkins University Press, Baltimore

Risch N (1990) Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46:229–241

Risch N, Zhang H (1996) Extreme discordant sib pairs for mapping quantitative trait loci in humans. Science 268: 1584–1589

Smith CAB (1959) Some comments on the statistical methods used in linkage investigations. Am J Hum Genet 11:289–304

Whittemore AS (1996) Genome scanning for linkage: an overview. Am J Hum Genet 59:519–528

Witte JS, Elston RC, Schork NS (1996) Genetic dissection of complex traits. Nat Genet 12:355–358